



RecordMe: A Smartphone Application for Experimental Collections of Large Amount of Data Respecting Volunteer's Privacy

David Blachon, François Portet, Laurent Besacier, Stéphan Tassart

► To cite this version:

David Blachon, François Portet, Laurent Besacier, Stéphan Tassart. RecordMe: A Smartphone Application for Experimental Collections of Large Amount of Data Respecting Volunteer's Privacy. UCAmI 2014, Dec 2014, Belfast, United Kingdom. pp.345–348. hal-01071484

HAL Id: hal-01071484

<https://hal.science/hal-01071484>

Submitted on 10 Oct 2014

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

RECORDME: a Smartphone Application for Experimental Collections of Large Amount of Data Respecting Volunteer's Privacy

David Blachon^{1,2}, François Portet¹, Laurent Besacier¹, Stéphan Tassart²

¹ Laboratoire d'Informatique de Grenoble
Univ. Grenoble Alpes, LIG, F-38000 Grenoble, France
CNRS, LIG, F-38000 Grenoble, France

41 rue Mathématiques, BP 53, 38041 Grenoble cedex9, France
{david.blachon, francois.portet, laurent.besacier}@imag.fr

² STMicroelectronics, 75669 Paris, FRANCE
stephan.tassart@st.com

Abstract. Since the spread of smartphones, researchers now have opportunities to collect more ecological data. However, despite the many advantages of existing databases (e.g., clean data, direct comparison), they may not suit all criteria for a particular experiment, resulting in an unavoidable tradeoff between the gain they provide and the lack of some labels or data sources. In this paper, we introduce RECORDME, an Android application ready to use for the research community. RECORDME allows to continuously record many different sensors and sources and provides a basic GUI for quick and easy settings. Also, a mark up interface is embedded for experiments that need it. Because of the high sensitivity of some data, RECORDME includes features for protecting volunteers' privacy and securing their data. RECORDME has already been successfully tested on different smartphones for 3 data collections.

Keywords: Mobile and sensor data acquisition, Context Sensing technology, Databases for data mining, Non-intrusive smart technology

1 Introduction

The recent spread of smartphones and tablets has opened many opportunities for the research community. Leveraging smartphones many resources, standard development frameworks, and the many potential volunteers, new sensing ways have emerged such as participatory sensing [1]. Another hot topic is the use of personal data collected through smartphones for tasks such activity recognition or context recognition [2]. In any of these cases, researchers must collect data to build systems and evaluate them.

Early studies have performed such sensing for collecting personal data such as Reality Mining [3], a campaign for collecting data through a particular smartphone provided to around 90 participants. More recently, Kiukkonen et al [4] collected data of almost 200 volunteers over several months. Both campaigns

provided useful databases (although they relied on a single brand of smartphones), containing volunteers' location, social interaction clues or accelerometer readings [4]. Existing databases offer many advantages such as handling tools, annotated data, provided evaluation methodology, easy and direct benchmarking.

However, such databases may not fit all the criteria, especially when it comes to use for a task a database that was acquired for another task. Hence, researchers have to trade research objectives off for the gain of time and money of using such databases. Moreover, existing databases are rarely provided with the tools that were used for their acquisition, which makes difficult to reproduce such acquisitions. In fact, to the best of our knowledge, no tool is available for the research community to collect personal data on smartphones.

In this paper, we introduce RECORDME, an Android application designed for the research community so as to collect data through volunteers' smartphones. RECORDME can continuously record data through around 20 different sources embedded on smartphones. A basic GUI allows an easy and quick use and can be configured. A mark up view is included in the GUI to allow volunteers to fill in personal annotations when experiments require it. Finally, because of the high sensitivity of some collected data, built-in features allow to protect volunteers' privacy and secure data. In Section 2, RECORDME is introduced while Section 3 reports first statistics and feedbacks of its use for data collections.

2 RECORDME: A Collecting Application for the Research Community

2.1 Design

RECORDME can record almost 20 different data sources at a time, both continuous or event-based, and provide either raw data or processed ones. Table 1 summarizes recorded sensors and processings if any.

This core feature is wrapped up with a basic GUI that allows volunteers a total control of the recording (see Figure 1). Indeed, she/he can select the sources to record, start and stop them at any time, and a notification informs her in real time about the status of the recording. If required, data transmission can be performed through a button on the main view that will start a wifi connection and transmit data through Secure Copy Protocol (SCP) to a predefined server.

Table 1. Summary of sensed sources

Type	Sensors	Data
Continuous	Accelerometer, Barometer, Ambient light, Magnetometer, Gyroscope	Raw 3-axis readings
Continuous	Audio	DFT coefficients magnitudes
Event	Proximity sensor, Headset, Battery Screen State, Communication Logs (Call, SMS, Data, Bluetooth, Wifi)	Status transitions (e.g., on/off) Service status (e.g., in service) Comm. status (e.g., ringing)
Event	Communication logs, application use	Hash function on character strings (ids, MAC addresses, app names)
Event	GPS, Wifi	Translation of coordinates

Because some experiments might require volunteers to report some notes or mark up, the GUI provides a view for such purpose. The markup template can be configured to provide predefined fields to fill in so as to make annotation quick and easy. However, volunteers can switch to text free areas for writing down their own annotations.

2.2 Privacy and Security

For ensuring privacy care and security of the data, RECORDME embeds built-in features. For instance, data older than 48 hours are deleted on a regular basis. Also, audio sensing is stopped when volunteers pass a call.

Specific processing are applied to sensitive data. Location coordinates from GPS or Wifi hotspots are translated then stored on the device so that it is not straightforward for non-experimenters to use them. For audio, RECORDME does not save raw signal but features computed on the fly. Coefficients magnitude of the Discrete Fourier Transform (DFT) are computed on 50% overlapping buffers of approximately 25 ms. Then, consecutive vectors are averaged on a window of approximately 200 ms. These settings guarantee that the raw signal cannot be reconstructed.

Also, character strings such as phone numbers from incoming calls or SMS, MAC addresses of nearby Bluetooth or Wifi devices, or cell tower ids are processed since they are sensitive. Since only correlation between the occurrence of those elements with other flows of information and with markups are of interest, a non reversible transformation by hash function is applied to protect privacy. A hash function associates a unique identifier to every input so that the distribution and correlation of the occurrences are unchanged.

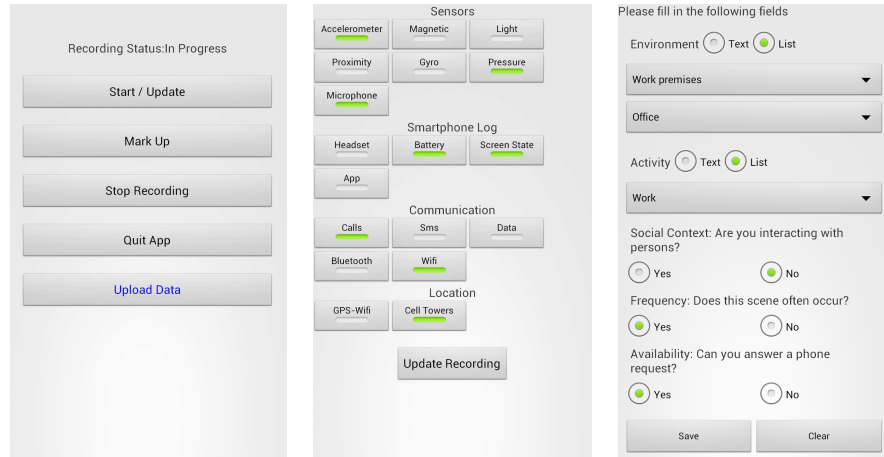


Fig. 1. Screenshots of RECORDME : from left to right, Home view (a), Source Selection view (b), and Markup view (c).

3 Conclusion

Three experiments have been run with RECORDME and are rich in feedbacks. First one is that RECORDME was stable on most tested smartphones. Only one smartphone had an original behaviour that consisted in stopping sensing when its screen turned off. We found that it was likely to be accountable for the constructor. Our workaround was the installation of a screen locker that kept the screen on then allowed a continuous recording. We also received volunteers' feedbacks about their experience with the application, especially that the marking up process was not convenient. Indeed, one collection campaign required volunteers to mark up just after the end of an activity. Some volunteers reported that they sometimes forgot: either they provided a delayed mark up or they just skipped it. Workarounds could be to ask volunteers to follow predefined scenarios or to orally record annotations.

Depending on experiments to perform and countries legislation, RECORDME can provide different security levels to protect volunteers' data. For instance, for an experiment that studies urban volunteers' motions and requires GPS coordinates collection, levels of security can be configured so that GPS coordinates are secured when stored on smartphones (by applying a translation for instance). Then, when transferred to the experimenters, coordinates are anonymized. Further processing of the coordinates is still possible by temporarily reversing the security processing (providing that volunteers agreed), yet the coordinates would remain anonymized. One of the data collections that used RECORDME performed in such a way and registered the protocol to the CNIL³.

So far, RECORDME has already been successfully installed and tested on more than 10 different smartphones from different brands (Acer, Google, Motorola, Samsung, Sony, Wiko) and Android OS versions 2.3 and 4.0. The three different data collections total almost 30 volunteers, located in different areas of France (mainly Grenoble, a South-Eastern town) and around the world (Budapest, Dublin, Singapour). 100 different recordings have been collected so far, resulting in a total of 430 hours.

References

1. J. A. Burke, D. Estrin, M. Hansen, A. Parker, N. Ramanathan, S. Reddy, and M. B. Srivastava. Participatory sensing. *World Sensor Web Workshop (at Sensys)*, 2006.
2. P. Chahuaara, F. Portet, and M. Vacher. Making context aware decision from uncertain information in a smart home : a markov logic network approach. In *Ambient Intelligence*, Lecture Notes in Computer Science, pages 78–93, Dublin, Ireland, 2013.
3. N. Eagle and A. Pentland. Reality mining: sensing complex social systems. *Personal and ubiquitous computing*, 10(4):255–268, 2006.
4. N. Kiukkonen, J. Blom, O. Dousse, D. Gatica-Perez, and J. Laurila. Towards rich mobile phone datasets: Lausanne data collection campaign. *Proc. ICPS, Berlin*, 2010.

³ Please visit its website for more details <http://www.cnil.fr/english/>. Our file reference is 0750532.